HP #59

Relative Value Scales for Physicians' Services

by

Jack Hadley, Ph.D.
Senior Research Associate
and
David A. Juba, M.P.A.
Research Associate

The Urban Institute
March 1983

The authors thank John Holahan [illegible]lip Held for their comments on a preliminary draft of this pape[illegible]thy Carlson and Ellen Pisarski for their excellent research assistance.

Support for this research was provided under contract HCFA-500-81-0053 with the Health Care Financing Administration, USDHHS. Opinions expressed are those of the authors and do not necessarily represent the views of The Urban Institute or its sponsors.

Public and private interest in using fee schedules to pay for physicians' services has been increasing as medical care costs continue to grow faster than expenses for other goods and services. A key element of any fee schedule is its underlying relative value scale (RVS) for physicians' services. A relative value scale is a cardinal ordering of the services provided by physicians. (Services must first be identified and defined by a procedure nomenclature, such as the AMA's CPT-4.) Cardinal ordering means that each service's numerical assignment provides both an ordinal ranking of all services and a measure of worth or value relative to a numeraire procedure. A relative value scale enables one to make statements like, "One service is worth ten times another, or is ten times more valuable, or is equivalent to ten units of another."

An RVS is not itself a fee schedule. To generate a fee schedule from the RVS, a dollar-per-unit conversion factor must be identified. For example, if an office visit has a scale value of 3 units and if the scale conversion factor is $10 per unit, the fee schedule value for the office visit is 3 x $10 = $30. The conversion factor for various categories of procedures may differ. For example, the conversion factor for medical procedures may be $10 per unit, while the factor for surgical procedures might be $15 per unit. Furthermore, separate conversion factors could be developed for particular procedures as well as for particular classes of procedures. The point is, an RVS measures the relative worth of each procedure, while a fee schedule, which need not preserve relative values, measures the absolute fee for each service.

An ideal RVS should reflect and balance the preferences and costs of patients, providers, and insurers. Patients want relative values to encourage physicians to consider financial and nonpecuniary factors (time, convenience, comfort) as well as medical value and efficacy in deciding how to provide

care. Physicians want relative values to take into account the costs of their time, their expenses for employees, equipment, supplies, and space, and the difficulty and riskiness of the cases they treat. In order to lower program costs, insurers want relative values to create incentives for both efficient provision of all services and the use of least-cost procedures when choosing among services of equal medical value.

Practical considerations, mainly the nature of available data bases and the costs of collecting new data, limit the construction of relative value scales to two basic types: those based on physicians' charges and those based on the amount of time physicians spend performing various services. Obviously, any relative value scale constructed from any data base can be altered or adjusted to either compensate for possible distortions inherent in the underlying data or to better reflect the preferences or objectives of the organization using the RVS. Nevertheless, it is important to understand the consequences for RVS construction of using different data bases or different construction methods. If issues of scale construction can be addressed and resolved independently, then evaluations of alternative scales and adjustments to scale values can be based on preferences and objectives without having to consider unnecessary complications.

This paper explores the characteristics of relative value scales constructed by alternative methods and using alternative data bases, four sets of charge data and one set of time data. We also compare the scales we construct to an RVS developed by a private, preferred provider organization, Mountain Medical Associates of Denver, Colorado. We examine the similarity of scales constructed from 1) different points on the distributions of procedure-specific charges (or times) from a particular data base, 2) different data bases, which contain information for different years and different geographic areas, and 3) all procedures grouped together and separate procedure classes.

The results of our analyses suggest that relative value scales constructed from physicians' charges are largely invariant with respect to construction method and data base. Pearson product-moment correlation coefficients and Spearman rank-order correlation coefficients among alternative scales were always positive, large in value, and significantly different from 0 at the 95 percent confidence level or better. The smallest correlation obtained was 0.82. The majority of correlation coefficients were greater than 0.90, with many larger than 0.95.

The correlations between time-based and charge-based relative values were also high for the various office, hospital, and home visits identified by CPT-4. Correlations were lower, however, for the limited number of other services (primarily surgical procedures) for which time data were available. Time spent by physicians in the performance of such procedures may not vary directly with either their investments in training or with the costs of complementary inputs (or other expenses) associated with these services. If so, one would expect to observe lower values of correlation coefficients between time-based and charge-based scales.

We reiterate that the objective of this study is to assess the consistency of procedures' values across alternatively constructed relative value scales. The study was not an attempt to evaluate scale values vis-a-vis some standard or norm. The analysis does not and cannot determine whether particular procedures are under- or over-valued. Other, outside information on medical efficacy, social costs, and affected parties' preferences must be brought to bear to make such judgments. The results do imply, however, that the choices of a data base and construction method are unlikely to alter the basic relative value scales.

METHODS

General Definition and Properties of a RVS

The relative value of the ith procedure or service is defined as $RVS_i = V_i/V_n$, where V_i and V_n are representative absolute values of the ith and the numeraire procedures, respectively. The representative absolute value of a procedure is a point selected from the distribution of charges (or times) for that procedure generated from observations of individual physicians' charges (or times). We compare up to four different points from each procedure's distribution of charges (or time): The mean and the median, which are common measures of central tendency, and the 75th and 90th percentiles, which are used by many insurers to screen individual physicians' billings for reasonableness.

Relative values are invariant with respect to scalar multiplication of the representative absolute values. For example, general price inflation which affects all procedures' charges equally will not affect relative values. Similarly, if all charges in large cities are higher than in small cities by a constant fraction, relative values will still be the same in both locales.

The choice of a procedure to be the numeraire is arbitrary, since the ordinal ranking of procedures is unaffected by the selection of the numeraire and cardinal values are preserved up to a multiplicative constant. In other words, for the ith procedure $RVS_{i1} = k_{12} \cdot RVS_{i2}$, where $k_{12} = V_{n2}/V_{n1}$, the ratio of the representative absolute values of the different numeraire procedures on the two scales. This means that it is not essential to identify separate numeraire's (thereby constructing separate scales) for surgical, medical, radiology, and pathology procedures.

Data Bases

Five sources of data were used in this study. The first is the Health Care Financing Administration's (HCFA) 1982 Prevailing Charge file.[1] It contains information on physicians' customary, unadjusted prevailing, and adjusted prevailing fees for 103 Medicare procedures in 238 Medicare localities, which are state or sub-state regions, across the country. (Medicare defines the customary fee as the median of each physician's charges for each service in an earlier time period, in this case calendar year 1980. The undjusted prevailing is the 75th percentile of the distribution of physicians' customary charges in the area. The adjusted prevailing fee results from the application of Medicare's Economix Index, which limits the rate of increase in unadjusted prevailings.) The fees reported to Medicare are based on charges by the medical specialties which provide each procedure most frequently. Most of the procedures are performed with high frequency.

The second and third data sources were constructed from all claims submitted to the California Medicare and Medicaid programs by a sample of over 7,000 physicians for services provided during the first calendar quarters of 1974 and 1978. The individual claims were obtained by The Urban Institute from the California insurance carriers which process claims for Medicare and Medicaid. All claims for procedures performed fewer than ten times in the aggregate were deleted, leaving 443 procedures for subsequent analysis. (These procedures accounted for more than 90 percent of all billings to Medicare and Medicaid.) Each physician's average billed charge and average amount reimbursed for each procedure were computed from the individual claims. The file also included the physicians' Medicare customary and prevailing charges. Since most physicians concentrate their billings among a relatively small number of procedures, the final analysis files contained

approximately 139,000 physician-procedure observations for each of the two time periods. The 1978 California claims file contained 83 procedures also included in the HCFA file and the 1974 California claims file contained 82 procedures also in the HCFA file.

The fourth data source was the Health Insurance Association of America's (HIAA) Surgical Prevailing Charge file. From surveys of 22 commercial insurance companies, HIAA obtained the mean, median, and other percentiles of charges for surgical procedures for which there were at least 5 claims in each of 250 geographic areas across the country. The period covered by this file was September 1977 through August 1978. Twenty-one of these surgical procedures were also included in HCFA's Prevailing Charge file.

The fifth data source was drawn from a national survey of medical and surgical specialties conducted by researchers at the University of Southern California.[2,3] Using log diaries sent to about 10,000 physicians between 1974 and 1976, the study collected information on approximately 250,000 patient-physician encounters. Each record included data on the:

- amount of time the physician spent with the patient
- encounter location (office, clinic, hospital, home, etc.)
- patient status (new or established)
- number of visits for this problem
- complexity (minimal, brief, limited, extended, comprehensive)
- severity (none, minor, acute, moderate, severe)
- urgency (none, deferrable, today, sooner, emergency)
- check-offs indicating performance (or ordering) of selected diagnostic or therapeutic proceduress

In our analyses, we employed such data from physicians in five specialties--general practice, internal medicine, general surgery, and pediatrics.

In addition to the five data bases used to construct relative value scales, we used an existing RVS developed by a physicians' practice

organization, Mountain Medical Affiliates Incorporated of Denver, Colorado.[4] As stated in the preface of their relative value study, "The first edition of the Mountain Medical Affiliates Relative Value Study (MMA-RVS) has been developed as a replacement for the 1971 Colorado RVS, which was found to be outdated. The most recent CURRENT PROCEDURAL TERMINOLOGY, Fourth Edition, is the source document utilized in determining procedure descriptions and their code numbers. The unit values that were developed by the Colorado Relative Value Committee in 1975 and 1976 were used as reference. In addition, many unit values in this study have been revised and added due to recent procedural and technological advances in the field of medicine." The MMA-RVS was developed over a two-year period by a series of specialty-specific committees charged with taking complexity, training, skill, outcome, and costs into account in setting relative values for procedures common to their specialty. We used this scale as a baseline against which to compare the constructed RVS.

These data bases differ in terms of geographic areas and specialties included, time periods, underlying data sources, and specific information reported. Comparing relative value scales constructed from such disparate sources provides tests of the sensitivity of relative values to the type of information used to compute them.

Analytic Techniques

The comparisons among alternative scales ask two basic questions. How similar is procedure ordering, and do procedures with large relative values on one scale also have large relative values on other scales? These questions were addressed by computing Spearman rank-order and Pearson product-moment correlation coefficients among scales. One advantage of correlation analysis is that coefficients are not affected by scalar differences among scales.

Coefficient values close to 1.0 indicate that the scales are substantially similar.

Ideally, one would like to compute correlation coefficients among all five data bases and the existing MMA-RVS simultaneously. However, use of the USC time data base was severely limited because procedures and visits reported on the log diaries were not identified by a common procedure nomenclature. (CPT-4 was used to identify procedures on each of the other data sources.) The only other, relatively large data base which reports time spent by physicians is the National Ambulatory Medical Care Survey.[5] It is smaller in size, is limited to office encounters, records less information than the USC data, and does not use a procedure coding terminology.

To partially compensate for this problem, we used the information on encounter location (office, hospital, home), patient status (new or established), number of visits for problem (first or subsequent), and complexity (brief, minimal, etc.) to assign CPT-4 visit codes (90000 through 90270) for fifteen different types of visits to all patient-physician encounters on the USC data file. Assigning CPT-4 codes for diagnostic or therapeutic procedures was more difficult because the descriptions of the diagnostic and therapeutic procedures on the log-diary reporting forms were much less precise and more inclusive than those used by CPT-4. As a result, only 13 nonvisit procedures could be assigned CPT-4 codes. For two of these procedures, heart catherization (CPT-4 code 93527) and proctosigmoidoscopy (CPT-4 code 45300), there were only 13 and 8 records respectively because these procedures were listed only for general surgeons (the only surgical specialists included in our data base).

Because of this difficulty with the USC data base, we analyzed the four charge-based data files and the MMA-RVS separately from the time-based RVS.

Comparisons between the time-based RVS and charge-based RVSs used the scale constructed from the HCFA Prevailing Charge data file, which, as will be shown, can be treated as representative of all of the charge-based RVSs analyzed.

To facilitate comparisons among the charge-based scales, we selected as the numeraire procedure needle puncture of bursa (CPT-4 code 20605), a procedure which was included in all four data bases and the MMA-RVS. As noted above, the designation of the numeraire procedure is arbitrary and does not affect the RVS in any substantive way.

Scales are identified by mnemonic names derived from the underlying data sources. The scale developed from Health Care Financing Administration data is labeled the HCFA scale; the two scales based upon The Urban Institute's California claims data are 78CAL and 74CAL, the fourth scale (developed from HIAA data) is called HIAA, and the Mountain Medical Associates RVS is identified as MMA.

RESULTS

Does the Choice of a Representative Absolute Value Affect the RVS?

The basic building block of an RVS is the charge or time used to represent a procedure's absolute value, V_i. For example, one could build RVSs from the mean, median, 75th or 90th percentiles of the distributions of either HCFA customary charges or prevailing charges across areas. How different are the resulting RVSs?

Twelve scales were calculated for the 103 procedures on the HCFA file and for the 443 procedures on the 1978 California file. The scales are based on 4 different points (mean, median, 75th percentile, and 90th percentile) from the distributions of 3 different types of charges (the adjusted prevailing, the

customary, and the unadjusted prevailing on the HCFA file; the average amount billed, the average amount reimbursed, and the customary on the 1978 California file). Four scales, based on the mean, median, 75th percentile, and 90th percentile of the distribution of average time per procedure across physicians, were computed from the USC file.

Pearson and Spearman correlation coefficients ranged from 0.96 to 0.99 among the scales generated from the HCFA data; from 0.95 to 0.99 on the 1978 California data scales; and, with one exception, from 0.90 to 0.99 among the time-based USC scales. (The one exception was a Pearson correlation value of 0.81 between the scales constructed from median and 90th percentiles of average time per procedure. This was due primarily to some very large time values for some of the procedures reported with low frequencies.)

From these results, we concluded that within a given data base, relative values are not sensitive to the selection of a particular type of charge (or time) or distribution point to be the representative absolute value of a procedure. This means, for example, that values on scales derived from billings data will not differ significantly from scale values constructed from reimbursement data. Similarly, one may choose the mean, median, or other point on charge distributions to serve as the representative charge for a procedure and not be concerned that the resulting scale will be affected by the choice. Consequently, any reasonably well constructed charge-based scale will serve as representative of the entire class of charge-based scales that one might construct from a given data base.

How Similar Are Scales Constructed from Different Data Bases?

Table 1 reports Pearson and Spearman correlations among five relative value scales under investigation. Only one scale is needed to represent each

Table 1

Correlations among RVSs Constructed from Different Data Bases (number of observations in parentheses)[1]

Pearson Product Moment Correlations[2]

	MMA	HIAA[3]	78CAL[4]	74CAL[4]
HCFA[5]	.978 (95)	.952 (21)	.998 (83)	.998 (82)
MMA		.972 (21)	.982 (77)	.979 (76)
HIAA			.999 (13)	.999 (13)
78CAL				.999 (82)

Spearman Rank-Order Correlations[2]

	MMA	HIAA[3]	78CAL[4]	74CAL[4]
HCFA[5]	.948 (95)	.978 (21)	.979 (83)	.975 (82)
MMA		.963 (21)	.960 (77)	.965 (76)
HIAA			.994 (13)	.996 (13)
78CAL				.994 (82)

Notes:
1. Limited to procedures common to all five RVSs.
2. All correlations significant at .0001 level.
3. Constructed from the 75th percentile of the distribution of customary charges across areas.
4. Constructed from the 75th percentile of the distribution of average billed charges across physicians.
5. Constructed from the 75th percentile of the distribution of median charges across areas.

of the data bases because of the high correlations among scales generated from the same data base. Only procedures common to all five scales are included.

All correlations are quite large, exceeding 0.94 in value. Correlations among RVSs calculated for four separate procedure groups, medicine, surgery, radiology, and pathology were somewhat lower on average, but still quite large. Forty-five of the 56 Pearson and Spearman correlation coefficients exceeded 0.90, and the remainder were between 0.82 and 0.89. All were statistically significantly different from zero at the five percent confidence level or better.

These results suggest that at least for these procedures, relative values are nearly identical across the five RVSs, in spite of the diversity of the underlying data sources. Two of the scales, 78CAL and 74CAL, came from similar data collected four years apart, a period over which absolute fees increased more than 45 percent. Two scales, HCFA and HIAA, are national in scope and based on area-wide median charges, while the 78CAL and 74CAL scales are based on individual physicians in a single state. The HIAA scale is constructed exclusively from charges to commercially-insured patients, while the 78CAL and 74CAL files reflect billings to Medicare and Medicaid, and the HCFA file comes form an amalgam of billings to public and private insurers, depending on how local carriers compute their fee screens. Finally, the MMA RVS represents a scale built from subjective professional evaluation of procedures' relative worth.

How Does a Time-Based Scale Compare to a Charge-Based Scale?

Constructing relative values for physicians' services from the amount of time a physician spends performing a procedure has considerable intuitive appeal. The physician's time is an obviously important component of the cost

of providing physicians' services. Time is relatively easy to measure, and the concept of using relative time to construct relative values is easy to understand. The time spent performing a procedure is unlikely to be directly affected by general price inflation or cost-of-living differences. In other words, time may be a more stable yardstick for constructing relative values than monetary values.

The key assumption in using time to construct relative values is that all other costs are approximately proportional to the amount of time the physician spends with a patient. This assumption may be tenable for some classes of procedures, such as visits and surgical procedures, where the physician's time is the primary input. For other procedures, especially radiology and pathology, the types and costs of equipment and support staff used may vary substantially. In addition, the physician's own investment in professional training varies significantly among the specialties associated with different procedures. For example, family practitioners, general internists, and pediatricians typically have three years of residency training compared to four for pathologists, five for general surgeons, and six to eight for surgical subspecialties. These differences in training investment would cause the value of time to vary among specialties. Thus, the basic unit of measure for constructing time-based relative values is in reality a variable, not a fixed yardstick.

Table 2 compares a time-based relative value scale with a charge-based relative value scale constructed from HCFA prevailing charge data. (Given the high correlations among the charge-based scales, selection of any other scale would not have changed the results.) Although the ordering of procedures on the two scales is fairly similar (Spearman correlation coefficient=0.89), the scale values are not as highly correlated as they were among the various

Table 2

Time-Based and Charge-Based
RV Scales, 25 Procedures

Procedures (CPT-4 Code)	Relative Value, Absolute Value (Scale Rank)	
	Time-Based	Charge-Based
1. Brief H.V. Estab. Pat. (90240)	0.68, 8.04 mins. (1.00)	0.96, $15.52 (3.00)
2. Minimal O.V. Estab. Pat. (90030)	0.72, 8.61 (2.00)	0.51, 8.26 (1.00)
3. Brief O.V. Estab. Pat. (90040)	0.77, 9.13 (3.00)	0.85, 13.72 (2.00)
4. Brief O.V. New Pat. (90000)	0.85, 10.08 (4.00)	1.32, 21.34 (8.00)
5. Limited H.V. Estab. Pat. (90250)	0.96, 11.44 (5.00)	1.25, 20.11 (5.00)
6. Limited O.V. Estab. Pat. (90050)	1.00, 11.87 (6.00)	1.00, 16.10 (4.00)
7. Brief H.V. New Pat. (90200)	1.11, 13.13 (7.00)	2.46, 39.69 15.00)
8. Chemotherapy (96030)	1.17, 13.88 (8.00)	1.30, 20.91 (6.00)
9. Limited O.V. New Pat. (90010)	1.23, 14.63 (9.00)	1.00, 27.38 (12.00)
10. ECG (93000)	1.36, 16.19 (10.00)	1.71, 27.68 (13.00)
11. Extended H.V. Estab. Pat. (90270)	1.40, 16.67 (11.00)	2.12, 34.18 (14.00)
12. Extended O.V. Estab. Pat. (90070)	1.50, 17.85 (12.00)	1.65, 26.71 (11.00)
13. Brief Home V. Estab. Pat. (90140)	1.52, 18.08 (13.00)	1.30, 20.92 (7.00)
14. Limited Home V. Estab. Pat. (90150)	1.69, 20.04 (14.00)	1.52, 24.54 (9.00)

Table 2 (Continued)

Time-Based and Charge-Based
RV Scales, 25 Procedures

	Time-Based	Charge-Based
15. Arthrocentesis (20610)	2.03, 24.12 (15.00)	1.59, 25.62 (10.00)
16. Comprehensive O.V. New Pat. (90020)	2.16, 25.68 (16.00)	3.14, 50.71 (19.00)
17. Comprehensive O.V. Estab. Pat. (90080)	2.27, 26.98 (17.00)	2.68, 43.25 (17.00)
18. Comprehensive H.V. New Pat. (90220)	3.14, 37.30 (18.00)	3.73, 60.18 (20.00)
19. Herniorrhaphy (49505)	3.52, 41.82 (19.00)	27.2, 439.38 (21.00)
20. Thoracentesis (32000)	3.71, 44.11 (20.00)	3.08, 49.75 (18.00)
21. Hysterectomy (58265)	4.73, 56.13 (21.00)	53.3, 861.21 (24.00)
22. Cholecystectomy (47600)	5.02, 59.62 (22.00)	43.7, 705.65 (23.00)
23. Proctosigmoidoscopy (45300)	5.22, 62.00[a] (23.00)	2.62, 42.34 (16.00)
24. Colon Resection (44140)	8.27, 98.25 (24.00)	60.20, 972.29 (25.00)
25. Heart Catherization (93527)	10.20, 120.92 (25.00)	33.00, 532.28 (22.00)
All Procedures		
Mean	2.65	10.16
Standard Deviation	2.35	17.93
Pearson Correlation	0.77	
Spearman Correlation	0.89	

Notes: a. Based on data for general surgeons only.

charge-based scales (Pearson correlation coefficient = 0.77). Dividing the procedures into visit and nonvisit groups reveals that this relatively low correlation is due primarily to disparities in the relative valuation of the nonvisit procedures. The Pearson correlation coefficients for the two groups are 0.88 for 15 visit procedures and 0.65 for the 10 nonvisit procedures.

In order to examine further how the two scales differ, we calculated the ratio of charge-based to time-based relative values for each of four groups of procedures, hospital visits, office visits, operations, and all other visits. If the two scales were substantially the same, i.e., all physicians' time were valued equally and all nonphysician costs were proportional to physician time, then the ratio of one set of scale values to the other would be identical, i.e., one scale would be a constant multiple of the other. In fact, the ratios of scale values in the four groups were substantially different. Charge-based relative values were 6.5 times greater than time-based relative values for surgical procedures, 50 percent greater for hospital visits, 20 percent greater for office visits, and 10 percent lower for the remaining procedures on the two scales.

It is apparent, then, that at least one of the two necessary conditions for equivalency between the two scales does not hold. There are several factors which could cause this. Two have already been mentioned, differences across physicians in the amounts of training and skill needed to provide different types of procedures, and differences in the kinds and costs of equipment and other personnel included in physicians' charges for the various procedure groups. A third factor may be variations in insurance coverage for different types of procedures. For example, a hospital visit is more likely to be covered by insurance because a hospitalized patient is more likely to

have satisfied the annual deductible than the average patient seen in the office.

Another possibly important factor influencing charge-based relative values is differences in interphysician competition by procedure type. Again, looking at relative values for office and hospital visits, competition among physicians and patients' price sensitivity are probably greater for office visits than for hospital visits. Once the patient is hospitalized, the physician is much closer to being a monopoly provider. As a result of these market forces, charges for time spent providing office visits would be expected to be lower than charges for time spent providing hospital visits, even if the physician's other costs may be higher in the office than in the hospital.

Sorting out the full effects of these factors will require further research. However, using other data from the USC time-based data file, we were able to provide some information pertinent to this issue by comparing several characteristics of office and hospital visits which have similar nominal (brief, limited, extended,and comprehensive) designations. The characteristics available from the log-diaries were severity of the illness, urgency of the visit, the primary specialty of the physician providing the visit, and whether the physician is board-certified. (All terms used to describe severity and complexity are taken directly from the log-diary reporting form.)

Table 3 shows that within each nominal visit category, hospital visits compared to office visits had higher proportions of more severe and more urgent cases and were more likely to be provided by physicians with specialty training and by board-certified physicians. A comparison of Limited Office and Hospital Visits (for Established Patients) is instructive. These two are

Table 3

Selected Characteristics of Office and Hospital Visits

Procedure (CPT-4 Code)	Characteristics						
	No. of Encounters	Severity[a]			Urgency[c]		
		None or Minor[b]	Moderate[b]	Severe[b]	None or Deferrable	Same Day	Sooner or Emergency
1. Brief Off. Vis., Estab. Patient (90040)	26,843	73.5%	24.0%	2.2%	49.1%	48.2%	2.3%
2. Brief Hosp. Vis., New Patient (90200)	3,949	38.1	46.6	14.9	19.6	71.0	8.4
3. Limited Off. Vis., Estab. Patient (90050)	29,736	53.2	42.4	4.3	42.6	53.4	3.5
4. Limited Hosp. Vis., Estab. Patient (90250)	7,905	19.3	56.1	24.6	12.3	82.5	4.7
5. Extended Off. Vis., Estab. Patient (90070)	3,595	19.8	61.1	18.8	39.0	53.0	7.2
6. Extended Hosp. Vis., Estab. Patient (90270)	4,571	4.9	52.2	42.8	9.3	72.1	18.1
7. Comp. Off. Vis., New Patient (90020)	1,485	36.65	48.1	14.7	45.8	33.3	20.4
8. Comp. Off. Vis., Estab. Patient (90080)	1,246	29.14	44.5	26.2	48.7	41.7	8.3
9. Comp. Hosp. Vis., New Patient (90220)	1,088	8.05	30.3	61.3	8.1	46.5	45.2
All Visits	80,418	50.4	39.2	10.2	38.3	55.8	5.4

Procedure (CPT-4 Code)	Characteristics				
	Primary Specialty of Physician[d]				Pct. Physicians Board-Certified
	Family or General Prac.	Gen. Surg.	Int. Med.	Ped.	
1. Brief Off. Vis., Estab. Patient (90040)	52.1%	5.7%	15.6%	23.9%	9.1%
2. Brief Hosp. Vis., New Patient (90200)	23.2	12.8	33.9	24.9	19.2
3. Lim. Off. Vis., Estab. Patient (90050)	38.8	3.5	19.2	35.7	12.1
4. Lim. Hosp. Vis., Estab. Patient (90250)	28.7	13.0	41.3	12.0	28.4
5. Ext. Off. Vis., Estab. Patient (90070)	43.5	6.2	31.4	12.8	20.5
6. Ext. Hosp. Vis., Estab. Patient (90270)	33.3	15.1	37.5	8.5	20.8
7. Comp. Off. Vis., New Patient (90020)	35.2	8.2	29.4	20.4	16.3
8. Comp. Off. Vis., Estab. Patient (90080)	31.8	2.8	47.6	12.9	33.5
9. Comp. Hosp. Vis., New Patient (90220)	18.8	6.3	50.9	14.3	26.7
All Visits	41.0	6.5	23.5	25.4	14.5

Notes: a. Percentage distribution of encounters in each procedure by severity.

b. Combines acute and chronic conditions.

c. Percentage distribution of encounters in each procedure by urgency.

d. Percentage distribution of primary specialties of physicians providing encounters.

the most frequently performed of the hospital and office visits reported in Table 3. Furthermore, the differences in characteristics between these two procedures are representative of the differences between all pairs of hospital and office visits reported there. The likelihood that hospital rather than office visits require more and/or more sophisticated medical treatment per episode is reflected in the relative frequency of "severe" cases in each. As the data of Table 3 reveal, almost 25 percent of all Limited Hospital Visits were so categorized, in contrast to only 4.3 percent of the Limited Office Visits. Furthermore, 82.5 percent of the hospital visits were adjudged by physicians to require "same-day" treatment; the corresponding figure for office visits was only 53.4 percent.

Even if hospital and office visits were of the same urgency and severity, and even if they consumed equal amounts of physician time, their relative values might differ due to differences in the training or skill levels of the attending physician. As shown in Table 3, 28.4 percent of the physicians performing Limited Hospital Visits in the USC data base were Board-Certified; this in contrast to the 12.1 percent of the physicians performing the Limited Office Visits. Also, proportionately fewer G.P.s (and more specialists) were among the physicians performing hospital, as contrasted with office visits. If one believes that physicians should charge more for treating more severe/complex cases or to recoup the costs of investing in more training, then at least some of the difference in the ratios of charge-based to time-based relative values for hospital and office visits seems to be appropriate.

DISCUSSION

This paper has examined alternative methods of constructing relative value scales for physicians' services. Developing a relative value scale is important because it is a key element in the construction of a fee schedule,

although it is not identical to a fee schedule. Fee schedules can be constructed from a relative value scale by applying a dollar multiplier or conversion factor that translates relative value units into monetary units. If the same dollar multiplier, for example $10 per relative value unit, is applied to all procedures, then relative fees would be identical to relative values. There is no inherent reason, however, why the same multiplier has to be used for all procedures, or for all medical specialties and geographic areas. If there are particular procedures (or medical specialties or geographic areas) for which public officials or private insurers wish to pay more, for whatever reason, then the use of different multipliers would accomplish this.

The results reported in this paper suggest that the underlying relative value scale, if constructed from data on physicians' charges, is quite robust with respect to the nature of the data used and the method of constructing relative values. The paper also compared a time-based relative value scale to a charge-based relative value scale. The correlation between the two was not as high as among the different charge-based scales. Two theoretical reasons for this are that simple average observed physician time per procedure does not capture very well variations in the value of different physicians' time or the costs of equipment and other personnel used in providing care. Further comparisons of several characteristics of office and hospital visits showed that hospital visits of a given nominal designation (brief, minimal, etc.) were more difficult, more urgent, and more likely to be provided by specialists than office visits of the same designation. Thus, both theory and evidence imply that relative values built from physicians' charges may be better indicators of the myriad factors that influence value than physicians' time per procedure.

Observed time could, in principle, be corrected to account for other factors. The one study of time-based relative values which attempted to adjust observed times to account for differences in overhead expenses, physicians' training investments, and procedure complexity found that for the small number of surgical pocedures examined, the agreement between the adjusted time and charge-based relative values "...is extremely close."[6] If cumbersome adjustments to time observations lead to a relative value scale essentially similar to one built from charges, then why not simply use charges to begin with?

To conclude that time is neither a sound nor practical basis for constructing relative values for physicians' services does not necessarily mean time should be irrelevant in determining physician payment. This paper was predicated on the assumption that the procedure, as identified by a procedure coding terminology, is the relevant unit of output for payment purposes. A separate issue is whether physicians should be paid on the basis of time regardless of the procedures performed. Time units (hours, days, weeks) would be the output measures, not procedures. Put in other terms, this is the salary vs. fee-for-service debate over how to pay physicians, which is not addressed by this paper. Furthermore, as is discussed below, policy makers and insurers may and can construct relative fees which create incentives for physicians to provide more time-intensive procedures than they might otherwise. Whether such incentives are warranted is also an issue beyond the scope of this paper.

Relative value scales built from charge data appear to be robust with respect to the data and method used and are relatively easy and inexpensive to construct. However, many argue that charges are distorted measures of absolute value because of the predominance of usual-customary-reasonable (UCR)

payment methods and differences in insurance coverage by class of procedure (operations are well insured but office visits and preventive care not not). Both UCR systems and insurance cause charges to be higher than they would be otherwise. But, if the inflationary effects of UCR and insurance are roughly the same for all procedures, then relative values will be unaffected.

UCR is unlikely to be distorting because it applies to all procedures. Insurance coverage is known to vary by procedure types, however. Future research needs to determine whether relative values for poorly insured procedures, such as office visits and preventive services, are similar when constructed from charges to a well insured population rather than a poorly insured population. If they are higher when based on charges to the well insured, then this would be consistent with the existence of insurance distortion.

This research contributes to the debate over fee schedules for physicians' services by suggesting that the issues of how to construct a relative value scale and whether different scales are needed for different purposes need not receive high priority. If the relative value scale is taken as given, then the major issues are the absolute levels of fees and the use of more than a single dollar multiplier to produce relative fees different from relative values. The first is essentially a physicians' earnings vs. insurers' costs issue, since relative fees would remain the same for any dollar multiplier, e.g., $5, $10, or $15 per relative value unit.

The second issue is more complex and has both equity and efficiency aspects. On equity grounds, one could argue that different multipliers should be used because of real differences in practice costs across geographic regions, different size cities, and/or medical specialties. Differential rewards or penalties should not be imposed on physicians because of factors

largely outside the implicit benefit/cost calculations that should be influenced by a fee schedule. The efficiency argument for different multipliers is that some pocedures are either over- or under-provided and relative fees need to be manipulated to correct these distortions. For example, if preventive services or so-called cognitive services lead to better health outcomes at equal or lower costs than curative or non-cognitive services, then preventive or cognitive services should have larger dollar multipliers than other services.

The other side of the coin in this debate, of course, is whether physicians respond to variations in relative fees in making various practice-related decisions. Various studies have shown that physicians do appear to be influenced by relative fees in deciding whether to treat Medicaid beneficiaries or accept assignment of Medicare benefits.[7,8,9] Other studies suggest that physicians' location choices are influenced by income opportunities in different areas, which presumably are affected by inter-area fee differences.[10,11] The evidence regarding the impact of fees and incomes on specialty choices is that there is at best a small influence,[11,12,13] although there is little or no recent research on this topic. Finally, there are no studies of the impact of relative fees on physicians' choices of specific medical procedures.

Constructing relative fee schedules which differ from the underlying relative value scale in order to influence physicians' behavior probably requires more research on how physicians respond to relative fees. In the short run, however, pressures to reduce spending for physicians' services and the rate of inflation of physicians' fees may lead to the adoption of fee schedules as cost cutting measures.

References

1. United States Department of Health and Human Services, Health Care Financing Administration: Medicare directory of prevailing charges, 1982.

2. Mendenall, R.C., R.A. Girard, and S. Abrahamson: A National Study of medical and surgical specialties. I. Background, purpose and methodology. JAMA 240:848-852, 1978.

3. Mendenhall, R.C., J.S. Lloyd, P.A. Repicky, et al.: A national study of medical and surgical specialties. II. Description of the survey instrument. JAMA 240:1160-1168, 1978.

4. Mountain medical affiliates relative value study. Denver, Colorado, Mountain Medical Affiliates, Inc., 1981.

5. Gagnon R., J. Delozier, T. McLemore: The national ambulatory medical care survey: 1979 summary. Vital and Health Statistics. Series 13-No. 66. (DHHS Publication No. (PHS) 82-1727). Public Health Service. Washington, D.C., Government Printing Office, 1982.

6. Hsiao, W.C., W.B. Stason: Toward developing a relative value scale for medical and surgical services. Health Care Financing Review Vol. 1, No. 2: 23-28, 1979.

7. Hadley, J.: Physician participation in medicaid: evidence from California. Health Services Research 14(4):266-80, 1979.

8. Sloan, F., J. Cromwell, J. Mitchell: Physician participation in state Medicaid programs. Journal of Human Resources 13(supp.):211, 1978.

9. Paringer, L.: Medicare assignment rates of physicians: their responses to changes in reimbursement policy. Health Care Financing Review Vol. 1, No. 3:75-90, 1980.

10. Hadley, J.: Financial incentives and physician distribution: evidence from Canada and the United States. Working Paper No. 1306-02-07, The Urban Institute, 1982.

11. Hadley, J.: Reimbursement, physicians' incomes, and physicians' specialty and location decisions, in Medical Education Financing. Jack Hadley, ed., New York: Prodist, 1980; 246-59.

12. Lee, R.H.: Scholarship programs and medical education financing, in Medical Education Financing, Jack Hadley, ed. New York: Prodist, 1980; 128-48.

13. Fruen, M.A., J. Hadley, S.P. Korper. Effects of financial incentives on physicians' specialty and location decisions. Health Policy and Education. 1980: 143-59.